

CLAIMS

1. A buffer element for a communication network, the buffer element comprising:  
a first buffer memory to store communication units corresponding to a first quality of  
5 service level;  
a second buffer memory to store communication units corresponding to a second  
quality of service level; and  
a buffer manager, coupled to the first buffer memory and the second buffer memory,  
to selectively store communication units in the first buffer and the second buffer based on a  
10 corresponding quality of service level of the communication units, and to retrieve  
communication units from the first buffer memory and the second buffer memory.
2. The buffer element of claim 1, wherein the buffer manager comprises:  
a sorter unit coupled to the first buffer memory and the second buffer memory to  
15 selectively store a communication unit in the first buffer or the second buffer based on a  
quality of service level of the communication unit.
3. The buffer element of claim 1, wherein the first buffer memory has a first depth, the  
second buffer memory has a second depth, and the buffer element further comprises:  
20 a depth adjuster to adjust the first depth and the second depth.
4. The buffer element of claim 3, wherein the depth adjuster comprises:  
means for iteratively searching possible depth assignments to determine the first depth  
and the second depth.  
25
5. The buffer element of claim 4, wherein the means for searching comprises:  
means for performing a steepest ascent hill climbing search.
6. The buffer element of claim 3, wherein the depth adjuster comprises:  
30 means for determining performance characteristics of the switch.

7. The buffer element of claim 1, wherein the first buffer memory and the second buffer memory are regions of memory in a contiguous random access memory device.

8. The buffer element of claim 1, wherein the communication units are ATM cells.

5

9. A switch for a communication network, the switch comprising:

a plurality of ports;

a first buffer memory coupled to one of the ports to store communication units corresponding to a first quality of service level; and

10 a second buffer memory coupled to the one of the ports to store communication units corresponding to a second quality of service level.

10. The switch of claim 9, further comprising:

15 a buffer manager, coupled to the first buffer memory and the second buffer memory, to selectively store communication units in the first buffer and the second buffer based on a corresponding quality of service level of the communication units, and to retrieve communication units from the first buffer memory and the second buffer memory.

11. The switch of claim 9, wherein:

20 the plurality of ports comprises a plurality of output ports that output communication units from the switch to the network; and

the first buffer memory and the second buffer memory are coupled to one of the plurality of output ports, to store communication units to be output to the one of the plurality of output ports.

25

12. The switch of claim 11, wherein:

each of the plurality of output ports has a respective first buffer memory and a respective second buffer memory to store communication units transmitted across the respective output port.

30

13. The switch of claim 12, wherein:

each of the plurality of output ports has a respective buffer manager to selectively store communication units in the respective first buffer and the respective second buffer based on a corresponding quality of service level of the communication units, and to retrieve  
5 communication units from the respective first buffer memory and the respective second buffer memory.

14. The switch of claim 9, wherein:

the plurality of ports comprises a plurality of input ports that receive communication  
10 units from the switch to the network; and

the first buffer memory and the second buffer memory are coupled to one of the plurality of input ports, to store communication units received on the one of the plurality of input ports.

15 15. The switch of claim 14, wherein:

each of the plurality of input ports has a respective first buffer memory and a respective second buffer memory to store communication units transmitted across the respective input port.

20 16. The switch of claim 15, wherein:

each of the plurality of input ports has a respective buffer manager to selectively store communication units in the respective first buffer and the respective second buffer based on a corresponding quality of service level of the communication unit, and to retrieve communication units from the respective first buffer memory and the respective second buffer  
25 memory.

17. The switch of claim 15, wherein the communication units are ATM cells.

18. A method buffering communication units in a communication network, the method  
30 comprising steps of:

assigning a queue depth for each of a plurality of queues, each queue being designated to store communication units of a predetermined quality of service level;

providing the plurality of queues, each queue having the corresponding assigned depth;

selecting one of the queues to receive a communication unit based on a quality of service level associated with the communication unit; and

5 storing the communication unit in the selected queue.

19. The method of claim 18, further comprising a step of adjusting the queue depths.

20. The method of claim 18, further comprising steps of:

10 monitoring a characteristic in the communication network; and

adjusting the assigned queue depths based on the monitored characteristic.

21. The method of claim 20, wherein the characteristic is selected from the group consisting of communication unit arrival rate for one of the quality of service levels, communication unit processing rate for one of the quality of service levels, communication unit loss rate for one of the quality of service levels and communication unit delay rate for one of the quality of service levels.

22. The method of claim 18, wherein each of the plurality of queues stores communication units for a single port in a communication network switch.

23. The method of claim 22, wherein the single port is an output port.

24. The method of claim 18, wherein the plurality of queues stores the communication units for each port of a switch in the communication network.

25. The method of claim 18, wherein the assigning step comprises a step of:  
determining a priority level for dropped communication units for each of the quality of service levels.

26. The method of claim 18, wherein the assigning step comprises a step of:  
assigning a priority level for communication unit delay for each of the quality of  
service levels.
- 5 27. The method of claim 18, wherein the assigning step comprises a step of:  
performing a search of possible depth assignments.
28. The method of claim 27, wherein the performing step comprises a step of:  
performing a steepest ascent hill climbing search.
- 10 29. The method of claim 18, wherein the communication units are ATM cells.
30. A method of selecting a communication unit, for transmission in a communication  
network that provides a plurality of quality of service levels, the communication unit being  
15 selected from a plurality of communication units stored in a buffer, the buffer including a  
plurality of queues, each queue corresponding to one of the quality of service levels, the  
method comprising steps of:  
identifying the queue with the highest corresponding quality of service level and  
which is not empty; and  
20 selecting the communication unit from the identified queue.
31. A method of storing a communication unit in a buffer, the communication unit having  
one of a plurality of quality of service levels, the buffer including a plurality of queues, each  
queue corresponding to one of the quality of service levels, the method comprising steps of:  
25 determining the quality of service level of the communication unit; and  
storing the communication unit in the queue having the corresponding quality of  
service level of the communication unit.
32. The method of claim 31, further comprising a step of:  
30 dropping the communication unit when the queue having the quality of service level  
of the communication unit is full.